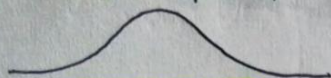


# MAXIMUM LIKELIHOOD

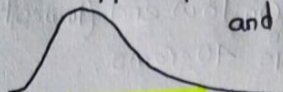
- The goal of maximum likelihood is to find the optimal way to fit a distribution to the data.
- There are lots of different distributions for different type of data and many more



NORMAL



Exponential



Gamma

- The reason you want to fit a distribution to your data is it can be easier to work and it is also more general - it applies to every experiment of the same type.

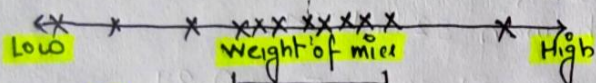
Example - let say we weighed a bunch of mice.

In this case, we think the weight might be normally distributed.

→ Avg (mouse weight)

Normally distributed means a lot of things

- i) We expect most of the measurements (mouse weight) to be close to mean (average)
- ii) We expect the measurement to be relatively symmetrical around the mean.
- iii) Normal distribution come in all shapes and sizes.



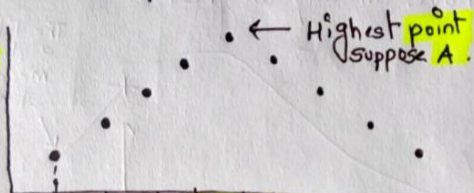
(i) Most of mice weigh close the avg

3 points

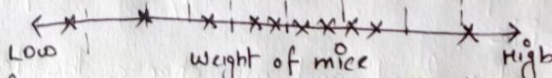
1 points

(ii) Although the measurement (3 & 1 points) are not perfectly symmetrical around mean, they are not crazy skewed to one side either.

Likelihood of observing the data



MEAN



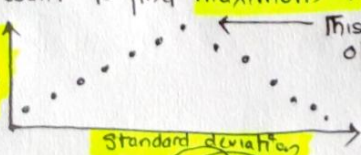
Point A from above graph

→ We want to locate "maximize the likelihood" of observing the weight are measured.

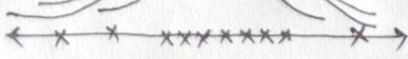
→ Thus it is the "maximum likelihood estimate for the mean".

If we want to find maximum likelihood of standard deviation →

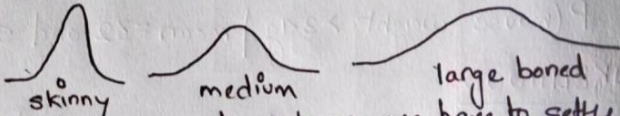
Likelihood of observing the data



Standard deviation

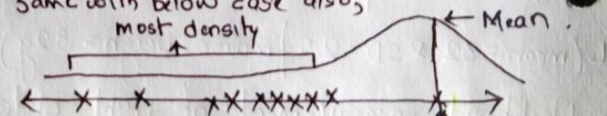


iv) Once we choose the shape, we have to settle out where will be the center thing i.e. most of the value should be near the average.



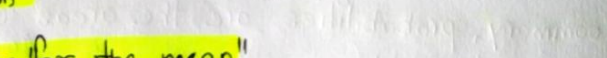
According to normal dist mean is here. Unfortunately most of the value are measured far from mean.

The probability/likelihood of observation is low. Same with below case also, most density



Best scenario, probability or likelihood of observing weight is high.

Plotting for every normal distribution



Now if someone say that maximum likelihood estimates for the mean or standard deviation or anything else we known that they found the value of mean or standard deviation that maximizes the likelihood that we observed the things we observed.